

The objective of this document is to recommend appropriate steps to be taken in the design of Access databases for the collection of long-term ecological monitoring data in the Denali National Park and Preserve. Specific attention is being placed on

- Providing a high level of documentation of the data set
- Quality Control
- Reporting results

I have reviewed the Denali National Park and Preserve Data Management Protocol and the sample data collected thus far. There are several general recommendations I have before I discuss specific steps.

Consistency. It is important that naming conventions, user interfaces, and data storage be as consistent as possible across the databases. For example, in one database it refers to the watershed where the sampling took place as “Drainage” and another database referred to it as “Location”. For ease of readability in the future, it’s best that the field names for comparable data are the same.

Documentation. This is essential so that anyone reviewing the data in the future can quickly identify the data set. There are places for internal documentation, such as the description for each field in the Table Design View. Each field should have a short description, with a sample entry, in the description column. Any programming code that may be written to support these databases should be profusely documented. Specific documentation practices can be found in the Leszynski Naming Conventions, or you can modify them to your needs. Lookup tables also make the data easier to read and interpret. An example in the sample databases that I reviewed was the Species field. The data was entered as abbreviations, thus making them less readable to someone who has not used those abbreviations. External documentation should be made as well. All documentation of fields, their attributes, queries, relations, reports, and modules should be made and printed out to be stored for later use. The Access Documentor can generate much of this for you (Tools|Analyze|Documenter).

Y2K Compliance. I highly suggest the use of a 4-digit date for date information. This should be done in the field in the table (set the format property to mm/dd/yyyy) and in data entry forms and reports. Note that if you set this format in the table before creating forms and reports with the Access Wizards, the Wizards will put the same format for the field on the form or report.

Key Fields. The identification of appropriate key fields is important. Primary Key fields uniquely identify each record in a table and their proper selection and creation is crucial for performance and data integrity.

Normalized Data. I suggest that the databases be normalized to at least the 3rd Normal Form and to the 5th Normal Form if it does not adversely affect performance and functionality. Normalizing in this way allows for the most flexibility and best performance.

- **First Normal Form** – Each field contains the smallest meaningful value and the table does not contain repeating groups of fields. In the following example, the name field should be broken down into at least first name and last name fields so records can be sorted by either first or last name and records could be located by either. Also, there are repeating groups of data identified by the Proj X and Time Proj X columns.

Emp ID	Name	Dept	Dept Name	Proj 1	Time Proj 1	Proj 2	Time Proj 2
EN1-26	Sean O'Brien	TW	Technical Writing	30-452-T3	25.00%	30-457-T3	40.00%
EN1-33	Amy Guya	TW	Technical Writing	30-452-T3	50.00%	30-482-TC	35.00%
EN1-35	Steven Baranco	TW	Technical Writing	30-452-T3	25.00%	31-238-TC	80.00%

- **Second Normal Form** – This applies only to tables with a multiple-field primary key. Each non-key field should not be a fact about only a part of the primary key, but should relate to the

entire primary key. In the following example, the primary key is based on the Emp ID and Proj Number fields. The Time On Project field relates to both the Emp ID and Proj Number. The Dept Code and Dept Name fields, however, relate only to Emp ID.

Emp ID	Last Name	First Name	Dept Code	Dept Name	Proj Number	Time On Project
EN1-26	O'Brien	Sean	TW	Technical Writing	30-452-T3	25.00%
EN1-26	O'Brien	Sean	TW	Technical Writing	30-457-T3	40.00%
EN1-26	O'Brien	Sean	TW	Technical Writing	31-124-T3	30.00%
EN1-33	Guya	Amy	TW	Technical Writing	30-452-T3	50.00%

- **Third Normal Form** – This applies only to tables with a single-field primary key. Each non-key field should be a fact about the primary key field, not about a non-key field. In the following example, the primary key is the Emp ID field. If you look at each record, the Last Name and First Name are facts about the Emp ID. Notice the Dept Code and Dept Name are not. The Technical Writing department is not a unique fact about Sean O'Brien. It actually is a fact relating to several employees. In this case we would move the Department information to its own table.

Emp ID	Last Name	First Name	Dept Code	Dept Name
EN1-26	O'Brien	Sean	TW	Technical Writing
EN1-33	Guya	Amy	TW	Technical Writing
EN1-35	Baranco	Steven	TW	Technical Writing
EN1-36	Roslyn	Elizabeth	AC	Accounting

Validation. I recommend using the strength of Access' field level validation to insure the data entered are correct. This is by no means the only Quality Control procedure, as the Protocols outline others. But this is the most effective as bad data are not allowed to be entered.

Combining Data. It may be efficient to combine several or all of these databases into one. However, I cannot recommend either way at this point in time. Once the analysis of the information and the determination of the tables and fields for each database are completed, there will be a better picture of the data needed and a determination will be possible.

Denali Management Protocol. A few items in the Protocol document deserve comment:

On page 2, it speaks of a "Master File". My interpretation of this is that there will be a database that will contain historical data and that each year's information will be added to it after it has been validated. If this is the case, this could raise questions on the use of primary and foreign keys. This could effectively negate the use of the Autonumber, depending upon how the Master File is structured. This would affect the database and table design steps noted below.

On page 5, in the paragraph headed "Software", it mentions that "source code should be organized by run order sequence." Since Access is an event-driven system, this is not possible as the run order is determined by the user's actions at run time.

Single- or Multi-User Use. It appears that one and only one person will use the databases at a time. However, I do recommend that they be developed with multiple user use in mind, as this could likely occur sometime in the future. In Access development, there's not a lot to do to make it multi-user, but it could affect the design phase. For example, it may be more efficient in a multi-user environment to not integrate as many of the current databases into one single database. This is, necessarily, a judgment call to be made by USGS and/or the National Park Service.

Specific steps I recommend at this time are:

1. Determine the data needed.

2. Make a list of all information that needs to be stored.
3. Look at report requirements to ensure no information is missed.
4. Have all involved attend the Level I and Level II Microsoft Access 8.0 classes.
5. Determine Naming Standards and Conventions
6. Determine Tables on paper
 - Main Tables
 - LookUp Tables
7. Determine Fields on paper
 - Common fields across databases
 - Names
 - Types
 - Sizes
 - Validation (Ranges, etc.)
 - Required Fields
8. Perform Normalization on paper
9. Have consultant review work thus far and make recommendations
10. Create databases and tables
11. Implement referential integrity
12. Have all involved attend the Level III, IV, and V Microsoft Access 8.0 classes.
13. Determine a standard interface for data entry
14. Create forms
15. Test work so far
16. Have consultant review work thus far and make recommendations
17. Create queries to support reports
18. Create reports
19. Test work so far
20. Have consultant review work thus far and make recommendations
21. Review documentation made during the entire process and supplement it as required.